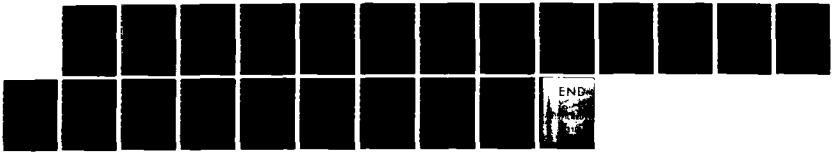
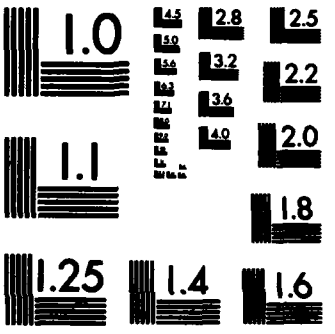
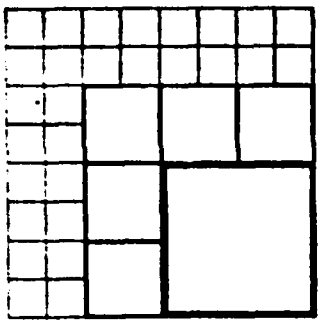
AD-A132 574 ADAPTIVE SPECTRAL SMOOTHING(U) MASSACHUSETTS INST OF TECH CAMBRIDGE STATISTICS CENTER F T PALMER AUG 83 TR-30 N00014-75-C-0555 1/1

UNCLASSIFIED F/G 12/1 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A



STATISTICS CENTER

Massachusetts Institute of Technology

77 Massachusetts Avenue, Rm. E40-111, Cambridge, Massachusetts 02139 (617) 253-8722

ADAPTIVE SPECTRAL SMOOTHING

BY

FERNANDO TUSELL PALMER

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

TECHNICAL REPORT NO. ONR 30

AUGUST 1983

PREPARED UNDER CONTRACT

N00014-75-C-0555 (NR-609-001)

FOR THE OFFICE OF NAVAL RESEARCH

ABSTRACT

A nonparametric, adaptive smoother for the periodogram is proposed, with a view to its application in linear prediction. Its performance is compared to several other existing methods, parametric and nonparametric, adaptive and non adaptive, of spectral estimation. The comparison is made in terms of prediction error.

AMS 1980 Subject Classifications: Primary, 62M
Secondary, 60G

Keywords and phrases: adaptive spectral estimation, linear prediction, spectral smoothing.

1. Introduction.

Spectral estimation has long been a subject of investigation. Since the early use of the periodogram by SCHUSTER(1898) a considerable effort has been directed to the improvement of spectral estimation. Major milestones were the adoption of periodogram smoothing and covariance weighting methods (see for example BARTLETT(1948) and BLACKMAN-TUKEY(1958)). These two essentially equivalent procedures dominated the spectral estimation scene for over two decades.

In the late sixties and early seventies a new class of parametric estimates was proposed (see AKAIKE(1969)). The method simply fits an autoregressive model $A(B)x(t) = e(t)$ and uses the estimated coefficients to compute an estimate $A(\omega)$ of the squared transfer function.

The spectral estimate is then computed as $\hat{f}(\omega) = (\hat{\sigma}^2/2\pi)\,|\hat{A}(\omega)|^{-2}$. The advantage of the method is the availability of criteria to truncate the autoregressive operator (such as FPE, CAT, AIC, etc.) which lend themselves to easy automatic implementation. This introduces some objectivity in spectral estimation.

In section 2 we present a procedure which tackles the same problem in the frequency domain. It is similar in spirit to the OSS method proposed by WAHBA(1980), although the implementation de-

parts substantially from Wahba's.

Section 3 describes Wahba's OSS and a third method for periodogram smoothing using a cross-validation idea proposed by STUETZLE(1982). In section 4 we report the results of a simulation. Comparisons are made in terms of square error of prediction. Section 5 contains some conclusions.

2. The CES method.

If we restrict our attention to the frequency domain, clearly the decision to be made is how and how much to smooth the periodogram $I(\omega_j)$. Once this decision has been made and $\hat{f}(\omega_j)$ obtained, factoring the spectrum estimate to get the AR coefficients is a purely numerical routine.

Under the conditions in section (1.1) and normality, asymptotically (BRILLINGER(1981)),

$$2[I(\omega_j)/f(\omega_j)] \sim \chi^2(2)$$

for $\omega_j = 2\pi j/N$ and $j = -N/2+1, \ldots, -1, 1, \ldots, N/2$, and:

$$I(\omega_j)/f(\omega_j) \sim \chi^2(1) \qquad \text{for} \quad j=0 \text{ or } N/2$$

Ignoring for the time being $\omega_j = 0$ or π and taking the (natural)

log of $I(\omega_j)$ we get:

$$\ln I(\omega_j) = \ln f(\omega_j) + \underbrace{\ln U_j - \ln 2}_{\xi_j} \qquad (2.1)$$

where U_j is a random variable distributed as $\chi^2(2)$ (except for $j=0$ or $N/2$).

We can find the mean and variance of $\xi_j = (\ln U_j - \ln 2)$ following DAVIS-JONES(1968) (also WAHBA(1980)). The characteristic function of $\ln U_j$ is;

$$\Psi(\theta) = E[e^{i\theta \ln U_j}]$$

$$= E[U_j^{i\theta}]$$

In general, if $X = U_j$ follows a gamma distribution with density,

$$g(x) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} e^{-x/\beta}$$

the cumulant generating function is:

$$k(\theta) = \ln \Psi(\theta) = \ln \int_0^{\infty} \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{i\theta} e^{-x/\beta} x^{\alpha-1} dx$$

$$= i\theta \ln \beta + \ln \Gamma(\alpha + i\theta) - \ln \Gamma(\alpha)$$

Differentiating with respect to $i\theta$ and setting $\theta = 0$ we get the first two cumulants:

$$K_1 = \ln \beta + A(\alpha)$$
$$K_2 = A'(\alpha)$$

The function $A(z) = (d/dz)\,\Gamma(z)$ is the psi (or digamma) function. Both $A(z)$ and its derivative can be approximated using asymptotic expansions (see DAVID-JONES(1968)). In our case, since $X = U_j$ (for $j \neq 0$ or $N/2$) is a chi square with two degrees of freedom, we have $\alpha = 1$, $\beta = 2$, and:

$$E[\ln U_j] = \ln 2 + A(1) = \ln 2 - \gamma$$
$$\text{var}[\ln U_j] = A'(1) = \pi^2/6$$

where $\gamma = .57721$ is the Euler constant. Going back to (2.1) we see that ξ_j ($j \neq 0$ or $N/2$) is white noise with mean γ and variance $\pi^2/6$.

We can heuristically regard the problem of recovering $\ln f(\omega_j)$ from $\ln I(\omega_j)$ as the problem of recovering a signal corrupted by noise. It seems reasonable therefore to apply known results in classical linear filtering theory. We can filter $\ln I(\omega_j)$ with a filter having transfer function (see YAGLOM(1963)),

$$c_{ff}(\tau) \,/\, c_{II}(\tau) \qquad (2.2)$$

where $c_{II}(\tau)$ denotes the spectrum of $\ln I(\omega_j)$ -considered as a stochastic process indexed over ω_j- and $c_{ff}(\tau)$ is the square of the Fourier transform of $\ln f(\omega_j)$. If we denote by $c_{\xi\xi}(\tau)$ the spectrum of ξ_j, $c_{II}(\tau) = c_{ff}(\tau) + c_{\xi\xi}(\tau)$.

We can compute a sample analogue of (2.2) and obtain from it a filter with the desired response. We propose to use this filter for the smoothing of the log periodogram (bias corrected as explained before) in order to obtain $\widehat{\ln f(\omega_j)}$. Then $\widehat{\ln f(\omega_j)}$ is used to compute the Wiener-Kolmogorov predictor.

In the estimation of (2.2) we can make use of the fact that $c_{\xi\xi}(\tau)$ can be aproximated by $\pi/12$; $c_{ff}(\tau)$ is then estimated as $\hat{c}_{II}(\tau) - \pi/12$.

The extreme points $\omega_j = 0$ and $\omega_j = \pi$ are handled in a similar way, except that the bias correction is done by adding a different constant; WAHBA(1980) and DAVIS-JONES(1968) give details in a different context.

Note that $c(\tau)$ is nothing but what BOGGERT et al.(1963) called the "cepstrum". The idea above was partially motivated by the perception that if there are ripples or periodic oscillations in

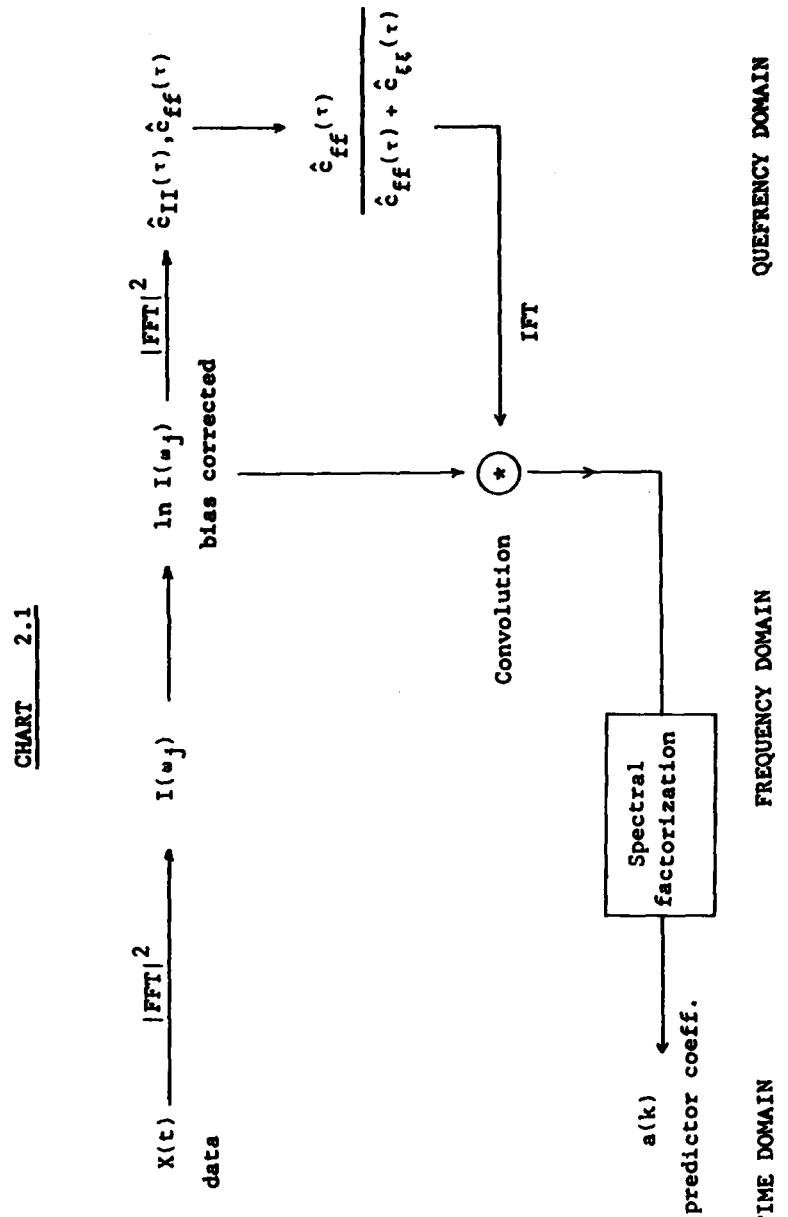
the "true" spectrum of a time series, the cepstrum is a powerful instrument to reveal their presence. Some empirical evidence is presented in BOGGERT et al.(1963), and we have run some simulations (not reported here) that seem to confirm this.

The following chart summarizes the operations to be performed in order to arrive at the CES predictor.

Note that we can either obtain the coefficients of a smoothing window by inverse Fourier transforming our estimate of (2.2) or damp the Fourier transform of ln $I(\omega_j)$ with (2.2) and then inverse Fourier transform. The last approach has been found less laborious, and on this ground has been used throughout.

The picture below raises some questions. For example, the goal was to construct an adaptive smoothing window to estimate $f(\omega_j)$ from $I(\omega_j)$. However, we have transferred the problem to the quefrency domain, where we again have to face the decision on how to smooth the squared Fourier transform of ln $I(\omega_j)$ in order to obtain $\hat{c}_{II}(\tau)$. It was thought that this smoothing in the quefrency domain (which we refer to as "pre-smoothing" in what follows) was a much less critical operation. The results of some simulations seemed to confirm this.

CHART 2.1



An additional problem is that $\hat{c}_{ff}(\tau)$ often happens to take negative values. This is something that could be expected since, if the spectrum of the time series is relatively smooth, for high quefrencies we have:

$$c_{ff}(\tau) \approx 0$$

and

$$c_{II}(\tau) = c_{ff}(\tau) + c_{\xi\xi}(\tau) \approx c_{\xi\xi}(\tau)$$

and since;

$$E[\hat{c}_{\xi\xi}(\tau)] = \pi^2/12$$

we can expect a substantial number of points at which,

$$\hat{c}_{ff}(\tau) - \pi^2/12 < 0$$

We have simply set $\hat{c}_{ff}(\tau) = 0$ whenever $\hat{c}_{ff}(\tau) = \hat{c}_{II}(\tau) - \pi^2/12$ happened to take a negative value.

There are several other issues that will be dealt with in chapter 4, where some advantages and weaknesses of the method are displayed along with some simulation results.

3. The SES and OSS methods.

The procedure outlined above is somewhat long and time consuming, although extremely simple conceptually -it reduces to the contemplation of ln $I(\omega_j)$ as signal plus noise and attempts to extract the signal.

STUETZLE(1982) suggested a method to adaptively smooth sequences which, although not derived from any obvious principle, has much to commend it because of its simplicity, and can be strongly defended on common sense grounds.

The idea is the following: smooth the given sequence using moving averages with diferent spans, and for each point chose the span minimizing some criterion locally. This allows for a sequence to be smoothed with many different moving averages, each one minimizing the criterion on a given region.

Given a sequence of length N in an array X(.), Stuetzle's smoothing method can be described as follows (all operations indexed by (i) should be performed for all i=1 to N):

1) Select several different spans, S1,S2,...,Sn.
2) Let the sequence to be smoothed be in the array X(.), of length N. Initialize an array RES(.) of length N setting RES(i) = BIG for all i, where BIG is a very

large number.

4) Pick span S1 from the list in 1).

5) Smooth X(.) with a moving average of the current span. When computing the moving average centered in X(i), omit the term X(i). Save the results in SMOAR(.).

6) Compute squared "cross validated residuals",

$$\text{PROV}(i) = [\, X(i) - \text{SMOAR}(i) \,]^{**2}$$

7) Smooth PROV(.) with a moving average of fixed span.

8) If PROV(i) < RES(i) then set RES(i) = PROV(i), set SPAN(i) = current span, and ANS(i) = SMOAR(i), else go to 9).

9) Repeat steps 5)-8) above for next span in list 1). If no more spans to try, go to 10).

10) Set SMOAR(i) equal to the moving average of length SPAN(i), including the term X(i).

Note that the last computation need not be started from scratch; it merely requires to modify the value already in ANS(i) including X(i).

Note also the smoothing operation in 7); this is done because it would be unwise to fix the span at point (i) based only on the best fit to that particular point (which might be an outlier). If some span is to be used at a given point, it makes sense to require it to be the best locally.

The method which we have labeled SES (Stuetzle's Smoothed Estimate) implements this idea. Implicit in step 6) above is,

$$E\ |f(\omega) - \tilde{f}(\omega)|^2$$

as a figure of merit. In our case, however, we are not interested in a good fit of $\tilde{f}(\omega)$ to the true $f(\omega)$ "per se". We rather want a spectral estimate that, when factored, leads to a minimum square error predictor. Since approximately (cf. WHITTLE(1952),

$$\frac{\sigma^2}{2\pi}\int_{-\pi}^{\pi}\frac{I(\omega)}{f(\omega)}\,d\omega \quad \approx \quad \frac{1}{N}\sum_{t=0}^{N-1}|\hat{e}(t)|^2$$

it would make sense to adopt as estimate $\tilde{f}(\omega)$ a smoothed periodogram minimizing,

$$\int_{-\pi}^{\pi}\frac{I(\omega)}{\tilde{f}(\omega)}\,d\omega$$

It immediately comes to mind that one could apply Stuetzle's method replacing step 6) by:

6') $\text{PROV}(i) = I(\omega_i)/\text{SMOAR}(i)$

where SMOAR(i) contains a moving average of periodogram ordinates with one of the spans in list 1). This has been tried, but it turned out that the use of the criterion $E|f(\omega) - \tilde{f}(\omega)|^2$ yielded better results, and has therefore been retained.

This method is appealing, because conceivably it would lead to locally different smoothers of the periodogram. Where the underlying spectral density is very steep, the span would most likely be chosen small; where the spectral density flattens, the chosen span would be larger.

Some simulation results displayed in section 4 illustrate the performance of this method.

The Wahba smoother.

Let $I(\omega_j)$ be the periodogram, and let $Y_j = \ln I(\omega_j) + C_j$, where C_j is an appropriate constant to correct the bias, as described in section 2. The estimate of $\ln f(\omega_j)$ proposed in WAHBA(1980) is:

$$\hat{g}_{N,m,\lambda}(\omega) = \sum_{j=-\frac{N}{2}+1}^{N/2} \frac{\tilde{g}_j}{(1 + \lambda(2\pi j)^{2m})} \exp(2\pi i j \omega) \qquad -\tfrac{1}{2} \leq \omega < \tfrac{1}{2} \tag{3.1}$$

where:

$$\tilde{g}_j = (1/N) \sum_{k=-\frac{N}{2}+1}^{N/2} Y_k \exp(-2\pi i k j/N) \tag{3.2}$$

In other words, Wahba's estimate is a smoothed periodogram, the smoothing being done by damping in the quefrency domain. This is achieved by multiplying the cepstrum by the function,

$$D_{\lambda,m}(\nu) = (1 + \lambda(2\pi\nu)^{2m})^{-1} \tag{3.3}$$

The significance of the method is that Wahba is able to show that if:

$$R_N(\lambda,m) = \int_{-\pi}^{\pi} [\hat{g}_{N,m,\lambda}(\omega) - g(\omega)]^2 d\omega \tag{3.4}$$

an approximately unbiased estimate of $R_N(\lambda,m)$ is given by;

$$\hat{R}_N(\lambda,m) = \sum_{\nu=-\frac{N}{2}+1}^{N/2} \left[|\tilde{g}_\nu|^2 - \frac{1}{N}\cdot\frac{\pi^2}{6}\right]\left[1 - \frac{1}{1 + \lambda(2\pi\nu)^{2m}}\right]^2 + \frac{1}{N}\cdot\frac{\pi^2}{6} \sum_{\nu=-\frac{N}{2}+1}^{N/2} \left[\frac{1}{1 + \lambda(2\pi\nu)^{2m}}\right]^2 \tag{3.5}$$

Once the cepstrum coefficients $\tilde{g}_\nu$ have been computed, $\hat{R}_N(\lambda,m)$ is fast enough to compute, and this enables the user to try many different values for λ and m in (3.3). One can then retain the values which minimize (3.5) and use them in (3.3).

The damping function (3.3) has not been chosen at random, but rather it has a number of interesting properties; it is the response of a Butterworth filter, and can be given a Bayesian and an

spline interpretation. More on this in WAHBA(1978).

4. Some Simulations.

A number of simulations was performed with different models, chosen to be as diverse as possible.

The sample size was fixed at N=512, and the innovations were chosen to be N(0,1). Normal noise was generated using the Marsaglia method, much faster than the Box-Muller algorithm and still exact.

The simulated processes were:

1. $X(t) = .3*X(t-3) + .7*X(t-7)e(2t-1) + e(2t)$
2. $X(t) = .9*X(t-24) + e(t)$
3. $X(7) = .9*X(t-12) + .7*e(t-1) + .5*e(t-2) - .9*e(t-3) + e(t)$
4. $X(t) = .5*e(t-1) - .06*e(t-2) + .45*e(t-15) + e(t)$
5. $X(t) = -.4*e(t-8) + .9*e(t-15) + e(t)$

The first model was chosen to be non-linear. Models 2, 3, and 5 are rather extreme examples of autoregressive and moving average mechanisms. Model 4 was previously used in a simulation study by BHANSALI(1973), and was adopted to compare results. Model 3 is a mixed model.

In all cases, 692 observations were generated. The first 100 were discarded, the next 512 used to compute the predictor, and the last 80 were used to estimate the mean square error of prediction one step ahead.

Four predictors were computed from each sample, according to the following specifications:

1. FPE : Orders up to 100 for the AR have been tried in all cases, and double precision used throughout.

2. WIENER-KOLMOG.: Four Daniell (flat) windows were used, averaging 19,15,11, and 7 periodogram ordinates.

3. CES : Pre-smoothing was performed by averaging 19 raw $c_{II}(\tau)$ ordinates.

4. SES : Ten different lengths were tried for the smoothing window: (N/20), (N/20)*.75, ..., (N/20)*.75**9, and a triangular rather than flat shape was used. Twenty five contiguous terms $[I(\omega) - \tilde{f}(\omega)]^2$ have been averaged to determine the span giving best local fit.

5. OSS (Wahba) : The parameter m (which controls the steepness of the damping function (3.3)) has been held constant at m=4, while λ (the parameter which really matters, since it controls the bandwidth) has been allowed to take values $10^{**}r$, with r going from -22 to -5 by .2 steps.

For each of the five models above 100 samples have been generated and all the e ht predictors tried (the conventionally smoothed Wiener-Kolmogorow has been computed using four different Daniell windows, as indicated above).

The results of the simulation for each of the predictors and each of the models are displayed in the tables below. The quantities whose sample means and standard deviations are shown are the mean square errors of prediction one step ahead. The graphical displays in A.2 and A.3 can be a bit misleading, because the pooled standard deviations are used -and it is apparent that in some cases the standard deviations are unlikely to be equal.

One would be tempted to conduct a one-way analysis of variance to test the equality of the means (and perhaps use Tukey's studentized range distribution to make pairwise comparisons). However, since all eight predictors have been computed using the same samples, one cannot safely admit independence across levels (if a

sample is atypical, or contains an abnormal number of outliers, one can expect a uniform improvement or degradation of the results obtained with all 8 predictors).

However we can use multivariate theory to test hypothesis on the mean vector (μ_1 , μ_2 , ... , μ_8). In particular, we are interested in testing hypothesis of the form:

$$a'\mu = (0 \ldots 1 \ldots -1 \ldots 0)\,\mu = 0$$

i.e., equality for the mean square error of different predictors. Multivariate theory (see, for example, MORRISON(1967), p. 121) states that the probability that all verify simultaneously,

$$a'\bar{x} - T^2_{\alpha;p,N-p}\sqrt{\frac{1}{N}a'Sa} < a'\mu < a'\bar{x} + T^2_{\alpha;p,N-p}\sqrt{\frac{1}{N}a'Sa}$$

where S = estimated cov. matrix, and $T^2_{\alpha;p,N-p}$ is the $100\alpha\%$ Hotelling T^2 quantile, is $1-\alpha$.

We have tested equality of the mean square error of prediction for all possible pairs of predictors. The results have been displayed below using the common convention of linking with solid lines the predictors whose mean square error of prediction is not significantly different . The lines are meant to link only the endpoints.

Since in some cases this would produce graphs with too many lines, we have used in some graphs dotted lines to link predictors whose MSE of prediction ARE significantly different at the 95% confidence level.

We wish now to comment briefly some salient features in the tables and graphs below.

Tables A.1 and B.1.

The model simulated is non linear. We can regard it as an autoregressive model with a random coefficient at lag 7. This can model time series with "bursts" occurring from time to time. Clearly, it is not the kind of model that a linear predictor can be successful in dealing with.

The results show poor performance of all predictors; the more intense the smoothing of the spectrum, the better the WK method with a Daniell window seems to be. Only when the smoothing is limited to 7 periodogram ordinates the WK predictor is significantly worse than most of the others. The SES is also significantly worse.

Table A.1

Model: $X(t) = .3*X(t-3) + .7*X(t-7)e(2t-1) + e(2t)$

LEVEL	N	MEAN	ST. DEV.
FPE	400	2.205	.884
WK19	400	2.238	.912
WK15	400	2.255	.923
WK11	400	2.285	.936
WK07	400	2.337	.953
CES	400	2.268	.906
SES	400	2.346	.959
OSS	400	2.220	.894

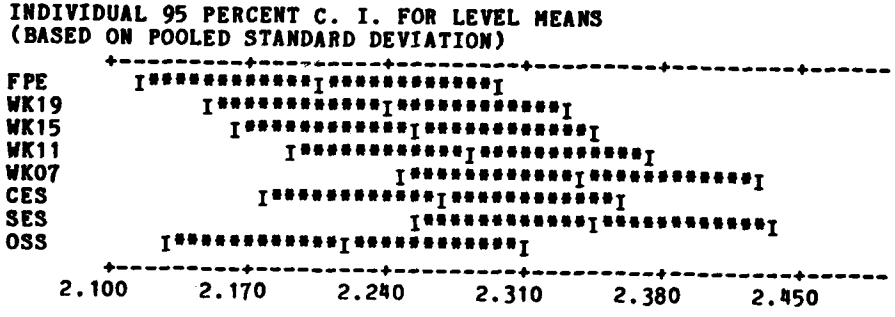


Table B.1

Model: $X(t) = .3*X(t-3) + .7*X(t-7)*e(2t-1) + e(2t)$

FPE	WK19	WK15	WK11	WK07	CES	SES	OSS

Table A.2 and B.2.

This is an extreme case of a spectrum with sharp peaks; the autoregressive operator has a root near the unit circle. The results displayed in tables A.2 and B.2 show how dramatic the degradation of the WK method can be, when the amount of smoothing applied to the periodogram is inadequate.

The autoregression fitted with the help of FPE shows an excellent performance, as one would expect; the generating mechanism is indeed an autoregression.

The second best is the CES predictor; this could also be expected, since the situation in which the spectrum is a function with strong periodicity is the one in which the choice of a customized smoothing window for the periodogram can give the greatest comparative advantage. This is so because the function that we are trying to recover can be filtered through a relatively narrow band filter, able to achieve a good separation of signal and noise.

Both OSS and SES give reasonable results. As for the conventional WK with a Daniell smoothing window, it is seen that the results can be disastrous if the smoothing is not well chosen.

Table A.2

Model: .9*X(t-24) + e(t)

LEVEL	N	MEAN	ST. DEV.
FPE	400	1.124	.226
WK19	400	4.803	1.322
WK15	400	3.462	.910
WK11	400	2.112	.504
WK07	400	1.660	.332
CES	400	1.329	.270
SES	400	1.633	.360
OSS	400	1.492	.333

INDIVIDUAL 95 PERCENT C. I. FOR LEVEL MEANS
(BASED ON POOLED STANDARD DEVIATION)

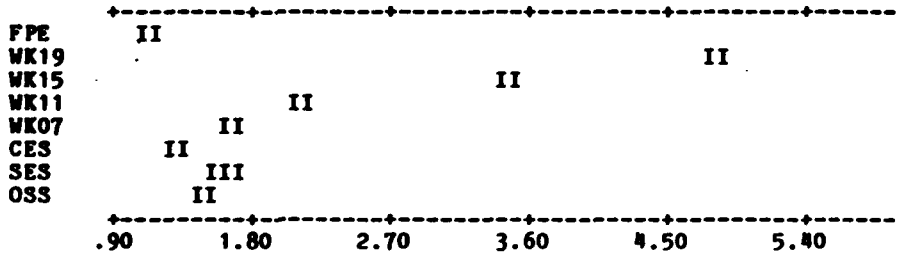


Table B.2

Model: X(t) = .9*X(t-24) + e(t)

FPE	WK19	WK15	WK11	WK07	CES	SES	OSS

Table A.3 and B.3.

This is a less extreme situation. The model is a mixed autoregressive-moving average (ARMA), with an AR operator whose roots are again near the unit circle. The autoregression truncated with FPE performs still uniformly better, while the three automatic methods for spectral smoothing give similar results. The use of a Daniell window averaging 7 periodogram ordinates is as good as the three automatic methods, but small increases in the amount of smoothing applied lead again to a poor showing of the WK predictor.

Table A.4 and B.4.

Now the sample is generated by a MA scheme, and the performance of the FPE shows a degradation, as one would expect. The three automatic smoothers (CES,SES, and OSS) give similar results. The WK with a Daniell window of variable length is again very sensitive to the amount of smoothing applied, and depending on how fortunate we are in the choice can give results which range from the best to the worst of all the considered.

When we look to the pairwise comparisons in table B.4 we see that the differences apparent in table A.4 are no significant, except for WK19. Still more, the confidence intervals for the linear compounds of the means are wide enough to rule out the possibility of increasing the number of replications to determine whether or not the perceived patterns are significant.

Tables A.5 and B.5.

All the comments for tables A.4 and B.4 seem relevant in this case, although the differences appear to be a bit more pronounced. However, the test for differences in the means fails again to reject equality of all pairs of means, except those involving WK19 and WK15, clearly over-smoothed.

Table A.3

Model: .9*X(t-12) +.7*e(t-1) +.5*e(t-2) -.9*e(t-3) + e(t)

LEVEL	N	MEAN	ST. DEV.
FPE	400	2.082	.370
WK19	400	3.729	.949
WK15	400	3.127	.714
WK11	400	2.606	.532
WK07	400	2.364	.436
CES	400	2.490	.510
SES	400	2.428	.460
OSS	400	2.460	.501

INDIVIDUAL 95 PERCENT C. I. FOR LEVEL MEANS
(BASED ON POOLED STANDARD DEVIATION)

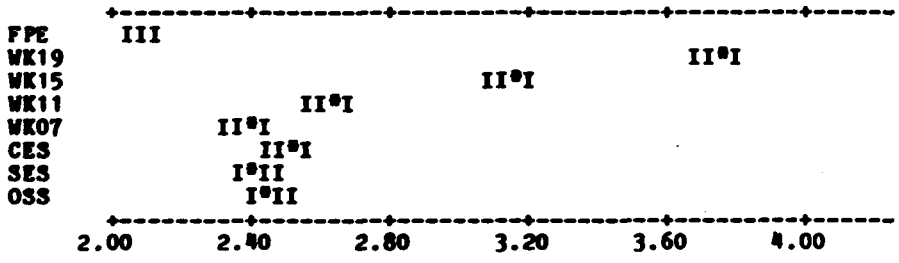


Table B.3

Model: X(t) = .9*X(t-12) + .7*e(t-1) + .5*e(t-2) - .9*e(t-3) + e(t)

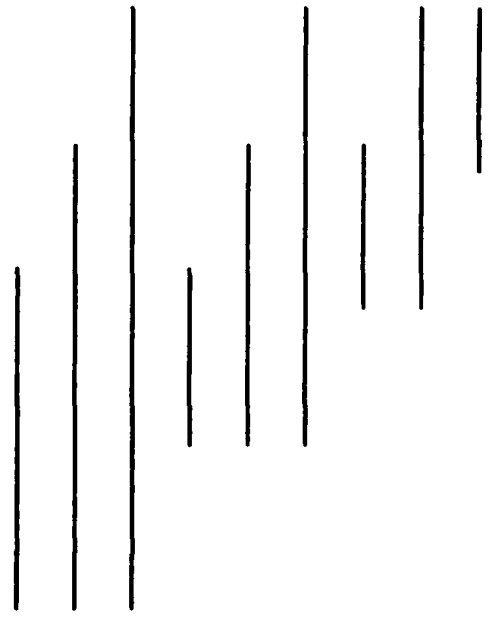


TABLE A.4

Model: X(t) = .5*e(t-1) - .06*e(t-2) + .45*e(t-15) + e(t)

LEVEL	N	MEAN	ST. DEV.
FPE	400	1.095	.190
WK19	400	1.113	.194
WK15	400	1.087	.190
WK11	400	1.071	.187
WK07	400	1.077	.187
CES	400	1.095	.194
SES	400	1.101	.192
OSS	400	1.094	.198

INDIVIDUAL 95 PERCENT C. I. FOR LEVEL MEANS
(BASED ON POOLED STANDARD DEVIATION)

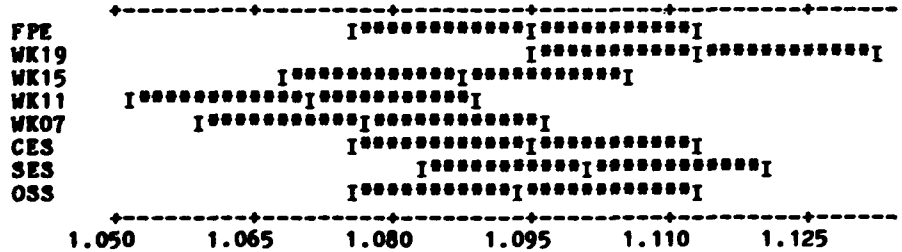


Table B.4

Model: X(t) = .5*e(t-1) - .06*e(t-2) + .45*e(t-15) + e(t)

FPE	WK19	WK15	WK11	WK07	CES	SES	OSS

Table A.5

Model: $-.4^{B}e(t-8) + .9^{B}e(t-15) + e(t)$

LEVEL	N	MEAN	ST. DEV.
FPE	400	1.407	.244
WK19	400	1.539	.270
WK15	400	1.454	.250
WK11	400	1.386	.238
WK07	400	1.354	.235
CES	400	1.419	.254
SES	400	1.410	.255
OSS	400	1.384	.249

INDIVIDUAL 95 PERCENT C. I. FOR LEVEL MEANS
(BASED ON POOLED STANDARD DEVIATION)

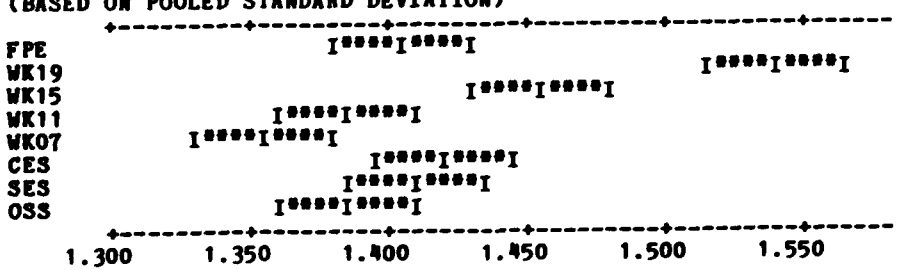
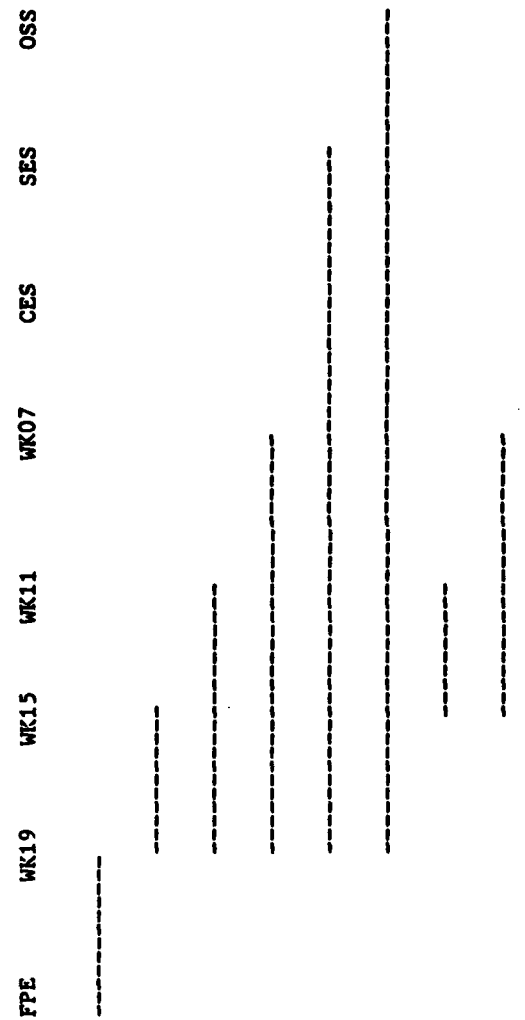


Table B.5

Model: X(t) = -.4*e(t-8) + .9*e(t-15) + e(t)



5. CONCLUSIONS.

All three methods tried for the automatic smoothing of the periodogram appear to give good results. CES seems to be the best choice in the case of spectra with strong regularities; otherwise, the OSS method proposed by Wahba appears to give reliable results in a wide variety of situations.

The results displayed also show that inadequate smoothing can have disastrous effects.

It should be borne in mind, however, that the conclusion above is made only on the basis of how good a linear predictor obtained by factoring the spectrum is. A spectral estimator should be judged not only according to its ability to produce "good" inputs to a spectral factorization routine, but also in the light of its ability to resolve peaks, its cost, and other features.

As far as linear prediction is concerned, however, if seems that a simple autoregression truncated using FPE is better than any of the other methods when the mechanism generating the sample is not far from autoregressive.

And, while a certain relative degradation is observed when the sample is generated according to a MA scheme, our simulation runs have failed to show FPE as significantly worse (at the 95% confi-

dence level) than any of the other methods.

Acknowledgements.

I am grateful to S. Ellis for his help and criticism. Whichever errors or inadequacies remain in this paper are solely my responsibility.

REFERENCES

AKAIKE,H.(1969) "Fitting autoregressive models for prediction", Ann.Inst.Stat.Math., vol. 21, p. 243-247.

AKAIKE,H.(1970) "Statistical Predictor Identification", Ann. Inst.Statist.Math.,vol. 22, 1970, p. 203-217

BARTLETT,M.S.(1948) "Smoothing periodograms from time series with continuous spectra", Nature, 161, 686-687.

BHANSALI, R.J.(1973) "A Simulation Study of the Wiener-Kolmogorov Predictor", Sankhya,A, vol 35, p. 357-376.

BLACKMAN,R.B.-TUKEY,J.B.(1958) "The measurement of power spectra from the point of view of communications engineering", Bell System Tech. J., 37, 183-282, 485-569.

BOGERT,B.P.,
TUKEY,J.W.,HEALY,M.J.R.(1963) "The Quefrency Analysis of Time Series", in Proc. Symp. Time Series Analysis, (ed. Rosenblatt), p. 209-243, J.Wiley & Sons, New York, 1963.

BRILLINGER, D.R.(1981) Time Series: Data Analysis and Theory, Holden-Day, Inc., 1981.

DAVIS,H.T.-JONES,R.H.(1968) "Estimation of the Innovation Variance of a Stationary Time Series", JASA,63,p.141-149.

JONES,R.H.(1976) "Estimation of the Innovation Variance of a Multivariate Stationary Time Series",JASA,71,386-388.

MORRISON,D.F.(1967) Multivariate Statistical Methods, McGraw-Hill, New York.

SCHUSTER, (1898) "On the investigation of hidden periodicities with application to a supposed 26 day period of metheorological phenomena", Terr. Magn., 3, 13-41.

STUETZLE, W.(1982) Class notes, MIT course 18.465, Spring 1982.

WAHBA,G.(1980) "Automatic Smoothing of the Log Periodogram", Periodogram",JASA, 75, p. 122-132.

WAHBA,G.(1981) "Data Based Optimal Smoothing of Orthogonal Series Density Estimates", Ann. of Stat., 9, p. 146-156.

WHITTLE, P.(1952) "Some results in Time Series Analysis", Skandinavisk Aktuarietidskrift, vol.35, p. 48-60.

YAGLOM (1962) An Introduction to the theory of Stationary Random Functions, Prentice-Hall, New Jersey, 1962.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER: 30	2. GOVT ACCESSION NO.: AD-A132 574	3. RECIPIENT'S CATALOG NUMBER:
4. TITLE (and Subtitle): Adaptive Spectral Smoothing		5. TYPE OF REPORT & PERIOD COVERED: Technical Report
		6. PERFORMING ORG. REPORT NUMBER:
7. AUTHOR(s): Fernando Tusell Palmer		8. CONTRACT OR GRANT NUMBER(s): N00014-75-C-0555
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Statistics Center, Massachusetts Institute of Technology, Cambridge, MA 02139		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: (NR-609-001)
11. CONTROLLING OFFICE NAME AND ADDRESS: Office of Naval Research, Statistics and Probability Code 436, Arlington, VA 22217		12. REPORT DATE: August 1983 13. NUMBER OF PAGES: 35
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report): Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:

16. DISTRIBUTION STATEMENT (of this Report)

This document has been approved for public release and sale; its distribution is unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Adaptive spectral estimation, linear prediction, spectral smoothing

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

See reverse side.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

A nonparametric, adaptive smoother for the periodogram is proposed, with a view to its application in linear prediction. Its performance is compared to several other existing methods, parametric and nonparametric, adaptive and non adaptive, of spectral estimation. The comparison is made in terms of prediction error.

END

FILMED

9-83

DTIC